A Reproduced Copy

OF

N83-33590

Reproduced for NASA

by the

NASA Scientific and Technical Information Facility

FFNo 672 Aug 65

ICASE

SPLINE APPROXIMATIONS FOR NONLINEAR HEREDITARY CONTROL SYSTEMS

(NASA-CR-172784) SPLINE APPROXIMATIONS FOR NONLINEAR HEREDITARY CONTROL SYSTEMS (NASA) 54 p HC A04/MF A01 CSCL 12A

N83-33590

Unclas
G3/64 10976

P. L. Daniel

Report No. 82-10
April 28, 1982

LIBRARY COPY

OCT 8 1982

LANGLEY RESEARCH CENTER
LIBRARY, NASA
HAMPTON, VIRGINIA

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the

UNIVERSITIES SPACE USRA RESEARCH ASSOCIATION

SPLINE APPROXIMATIONS FOR NONLINEAR HEREDITARY CONTROL SYSTEMS[†]

P. L. Daniel[*]
Southern Methodist University

ABSTRACT

A spline-based approximation scheme is discussed for optimal control problems governed by nonlinear nonautonomous delay differential equations. The approximating framework reduces the original control problem to a sequence of optimization problems governed by ordinary differential equations. Convergence proofs, which appeal directly to dissipative-type estimates for the underlying nonlinear operator, are given and numerical findings are summarized.

[*]Part of this research was completed while the author was a visitor at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA, operated under NASA Contracts No. NAS1-15810 and NAS1-16394.

[†]Work supported in part by the Air Force Office of Scientific Research under Contract AFOSR 76-3092D, in part by the National Science Foundation under Grant NSF-MCS 7905774-05, and in part by the U. S. Army Research Office under Contract ARO--DAAG29-79-C-0161.

§1 Introduction

The focus of our efforts here is to extend the results of Banks in [3] to spline-based approximation schemes for control systems governed by nonlinear nonautonomous functional differential equations (FDE). Our ideas are similar to [3] in spirit in that we approximate infinite-dimensional FDE systems by finite-dimensional ordinary differential equation (ODE) systems. (This approach is not new to the treatment of control problems; a fairly thorough summary of the literature may be found in [6].)

Our results improve upon the findings of [3] (as well as the linear FDE results in [10], [5], [6] and the nonlinear autonomous findings of [21]) in that we handle very general nonlinear nonautonomous FDE systems that may contain nonlinear discrete delay terms (i.e., terms involving x(t-r)); in [3], discrete delay terms could only appear in the linear autonomous part of the control system. Additionally, our approximations are based on linear splines which have generally demonstrated superior performance when compared in [10] and [7] to the averaging

approximations treated in [3]. Furthermore, we adopt a more direct theoretical approach than that of [3]. In the latter, the nonlinearity is treated as a perturbation of a linear system (which is formulated as a linear semigroup problem) and Trotter-Kato convergence results are invoked to demonstrate the convergence of approximating semigroups. Our ideas are in the spirit of [4], [8], and [9] where convergence is demonstrated through a direct application of dissipative-type estimates to the underlying nonlinear operators. This straightforward approach easily accommodates both nonlinear and nonautonomous systems, differing from such recent studies as the examination of linear non-autonomous FDE systems in [11] via an evolution operator analogue of the Trotter-Kato theorem, or the consideration in [19], [20] of nonlinear autonomous equations using a nonlinear Trotter-Kato type result.

Disadvantages of our treatment of FDE-governed control problems are that we require additional smoothness hypotheses on the right-hand side of the FDE and we must consider a restricted class of controls: Specifically, we will limit our attention to controllers with inertia, or those controls with restrictions on the rate at which they may be varied. Although such a restriction is not unreasonable in many applications (see the wind tunnel example in §4; information on relevant mechanical and economic systems may be found on pages 120-123 of [13]) it represents a limitation not found in the previously mentioned papers.

Our presentation is as follows: In section 2 we describe an optimal control problem for nonlinear FDE systems and present an equivalent Hilbert space formulation based on an abstract nonlinear evolution equation. Approximating spline subspaces are constructed in section 3, where approximate control problems are defined and convergence results are detailed. Finally, in section 4 we

present numerical findings for several examples where the spline approximations have been used to estimate the control and state variables.

Much of the notation is standard and in accordance with popular usage. Since most functions considered will be R^n-valued, we will designate by $H^p(a,b)$ and $L_p(a,b)$ the corresponding Sobolev and Lebesgue spaces of R^n-valued functions defined on (a,b); similarly, the space $C^m([a,b];R^n)$ of n-vector valued functions on $[a,b]$ with m continuous derivatives will be denoted by $C^m(a,b)$. For a square-integrable function $s \to x(s) \in R^n$, the notation x_t will designate the $L_2(-r,0)$ function given by $x_t(\theta) = x(t+\theta)$, $-r \leq \theta \leq 0$, where $r > 0$ is fixed throughout. Finally, we shall not distinguish between the column form of a vector and its transpose, and the same symbols $|\cdot|$ (and $\langle \cdot,\cdot \rangle$) will be used to denote any of several norms (and inner products) for R^n, L_2, and $Z = R^n \times L_2$ when the meaning is clear from the context.

§2 The Nonlinear FDE Control Problem

In the present section we describe the nonlinear nonautonomous FDE control problem that is the focus of this paper, and outline the properties that will be employed in subsequent sections. Our approach in §2.1 is to reformulate the FDE-governed control problem as an abstract problem on an infinite-dimensional state space. In Theorem 2.3, the result that concludes the section, we guarantee the existence of a control that minimizes the quadratic cost functional and generates a solution to the nonlinear FDE.

The control problem examined in this paper will be limited to a class of problems involving inertial controllers, i.e., those controls with restrictions on the rate at which they may be varied. The control problem considered here (referred to hereafter as *P*) may be formally stated as follows:

ORIGINAL PAGE IS
OF POOR QUALITY

Find an optimal control $\bar{u} \in U$, where U is defined by

$$U = \{u \in H^1((a,b);R^m) \mid |u(a)| \leq K\}$$

(K a fixed constant), that minimizes the quadratic cost functional J(u),

$$\begin{aligned} J(u) = {} & (x(b) - \kappa(0))Q_0(x(b) - \kappa(0)) \\ & + \int_{-r}^{0} (x_b(\theta) - \kappa(\theta))Q_1(x_b(\theta) - \kappa(\theta))d\theta \\ & + \int_{a}^{b} x(t)Q_2x(t)dt \\ & + \int_{a}^{b} \{u(t)Ru(t) + \dot{u}(t)S\dot{u}(t)\}dt. \end{aligned}$$

Here $(\kappa(0),\kappa)$ is a given "target" in $R^n \times L_2(-r,0)$, $Q_i \geq 0$ are $n \times n$ matrices ($i = 0, 1, 2$), R and S are $m \times m$ matrices with $R \geq 0$ and $S > 0$, (without loss of generality Q_i, R, and S may be assumed to be symmetric), and the pair (x,u) satisfies the vector nonlinear FDE

$$(2.1) \qquad \begin{cases} \dot{x}(t) = f(t,x(t),x_t,x(t-r_1), \ldots, x(t-r_\nu)) + B(t)u(t), \ a < t \leq b, \\ (x(a),x_a) = (\phi(0),\phi). \end{cases}$$

We shall assume that $B \in H^1((a,b);R^{n\times m})$, the initial data $(\phi(0),\phi)$ is such that $\phi \in H^1(-r,0)$, and the discrete delays satisfy

$$0 = r_0 < r_1 < \ldots < r_\nu = r;$$

in addition, f satisfies the following standing hypotheses.

ORIGINAL PAGE IS
OF POOR QUALITY

(H1) The mapping f satisfies a global Lipschitz condition on $R^n \times L_2(-r,0) \times R^{n\nu}$. That is, there exists $m_1 \in L_2((a,b);R)$, $m_1 > 0$, such that for all $(\xi,\psi,w_1,\ldots,w_\nu)$, $(\delta,\chi,y_1,\ldots,y_\nu) \in R^n \times L_2(-r,0) \times R^{n\nu}$,

$$|f(t,\xi,\psi,w_1,\ldots,w_\nu) - f(t,\delta,\chi,y_1,\ldots,y_\nu)| \le m_1(t)\{|\xi-\delta| + |\psi-\chi| + \sum_{i=1}^{\nu} |w_i-y_i|\}$$

for almost all $t\in[a,b]$.

(H2) The function f is differentiable on $[a,b]\times R^n \times L_2(-r,0) \times R^{n\nu}$.

(H3) The mapping $(\sigma,\psi) \to f_t(\sigma,\psi(0),\psi,\psi(-r_1),\ldots,\psi(-r_\nu))$ is continuous from $[a,b] \times C(-r,0)$ to R^n.

Remark 2.1 It is an immediate consequence of hypotheses (H1) that f satisfies an affine growth condition; that is, for x given in $L_2(a-r,b)$,

$$(2.2)\qquad |f(t,x(t),x_t,x(t-r_1),\ldots,x(t-r_\nu))|$$
$$\le m_1(t)\{|x(t)| + |x_t| + \sum_{i=1}^{\nu} |x(t-r_i)|\} + m_2(t),$$

where $m_2(t) = |f(t,0,\ldots,0)|$ is in $L_2((a,b);R)$ from (H2). Quite standard arguments may then be employed to demonstrate that $t \to f(t,x(t),x_t,x(t-r_1),\ldots, x(t-r_\nu))$ is in $L_1(a,b)$ and that the mapping depends only on the equivalence class of x. Therefore, since we shall only consider (2.1) in its equivalent integral form, there will be no difficulty associated with the point evaluations of x appearing in f.

Remark 2.2 Assumption (H3), which is required in the proofs of Lemma 3.2 and Lemma 3.4, may actually be replaced by a weaker hypothesis:

(H3') For $x \in E$, E bounded in $C(a-r,b)$, the mapping $\sigma \to f_t(\sigma,x(\sigma),x_\sigma,x(\sigma-r_1),\ldots, x(\sigma-r_\nu))$ is dominated by a $L_2(a,b)$ function, uniformly in $x \in E$.

ORIGINAL PAGE IS
OF POOR QUALITY

Remark 2.3 It is not difficult to prove that the set $\mathcal{U}$ of admissible controls is convex and closed in $H^1((a,b);R^m)$. In addition, we remark that although $u \in \mathcal{U}$ has been defined without constraints on the $L_2(a,b)$ norm of its derivative, it is clear that any optimal control $\bar{u}$ ($J(\bar{u}) < \infty$) must satisfy $|\dot{\bar{u}}| < \infty$.

Before we consider the problem of locating an optimal control $\bar{u} \in \mathcal{U}$, we shall first establish the existence, uniqueness, and continuity properties of solutions to (2.1) for any given u in $\mathcal{U}$. In the theorem that follows we state these results without proof, since the arguments required involve quite standard applications of uniform contraction principles ([18], p. 7), hypothesis (H1), and the growth condition (2.2) (see §1.3 of [14] for details).

Theorem 2.1 There exists a unique (continuous) solution x to (2.1) on the interval $[a-r,b]$ which depends continuously on $\{\phi(0),\phi,u\} \in R^n \times H^1(-r,0) \times H^1((a,b);R^m)$ in the $R^n \times L_2(-r,0) \times L_2((a,b);R^m)$ topology.

§2.1 An Abstract Reformulation of the Control Problem

We next establish an equivalence between the nonlinear nonautonomous FDE (2.1) and an abstract evolution equation (AEE) in an infinite-dimensional state variable. The abstract framework established in the present section is similar in spirit to [10] and [3] and is well-established in the FDE literature (see, for example, [6] and the references therein).

The state space will be defined to be the infinite-dimensional Hilbert space $Z = R^n \times L_2(-r,0)$ with norm $|\cdot|$ induced by the inner product $\langle(\xi,\psi),(\delta,\chi)\rangle \equiv \xi\delta + \int_{-r}^{0} \psi\chi$. For (t,ξ,ψ) in $[a,b] \times R^n \times C(-r,0)$ we define $F(t,\xi,\psi) = f(t,\xi,\psi,\psi(-r_1),\ldots,\psi(-r_\nu))$ and $A(t) : W \to Z$ by

$$A(t)(\psi(0),\psi) = (F(t,\psi(0),\psi), \mathcal{D}\psi) \tag{2.3}$$

where $W = \{(\psi(0),\psi) \in Z \mid \psi \in H^1(-r,0)\}$, and $\mathcal{D}\psi$ denotes the $L_2(-r,0)$ function that is the derivative of ψ.

ORIGINAL PAGE IS
OF POOR QUALITY

Details of the reformulation of (2.1) as an equivalent evolution equation on Z may be found in [8], where the same FDE is examined in the context of a parameter estimation problem (where coefficient parameters as well as delays $r_1,\dots,r_\nu$ and initial data $(\phi(0),\phi)$ are unknown). The results in section 2 of [8] may be readily extended to the FDE control system under consideration here; we state those findings in Lemma 2.1 and Theorem 2.2. The first of these results describes how to construct an equivalent topology for Z so that the nonlinear operator A satisfies a dissipative-type inequality. The lemma, a nonlinear analogue of similar estimates found in [6] and [10], greatly simplifies our calculations and is the foundation for our theoretical development without the use of semigroups.

Lemma 2.1 Let $y = (\xi,\psi)$ and $z = (\delta,\chi)$ be given in Z and define a new weighted inner product on Z by $\langle y,z\rangle_\rho = \xi\delta + \int_{-r}^{0} \psi(\theta)\chi(\theta)\rho(\theta)d\theta$ where ρ is given on $[-r,0]$ by

$$\rho(\theta) = \begin{cases} 1, & \theta \in [-r_\nu,-r_{\nu-1}) \\ 2, & \theta \in [-r_{\nu-1},-r_{\nu-2}) \\ \vdots & \vdots \\ \nu, & \theta \in [-r_1,0]. \end{cases} \tag{2.4}$$

Then $\langle A(t)y - A(t)z, y-z\rangle_\rho \le \omega(t)|y-z|_\rho^2$ for all $y,z \in W$ and almost all $t \in [a,b]$. The function $\omega > 0$ is in $L_1((a,b);R)$ and is given by $\omega(t) = \frac{3}{2}m_1(t) + \frac{\nu}{2} + \frac{\nu}{2}m_1^2(t)$.

The equivalence of the FDE (2.1) to an evolution equation in $z(t) \in Z$ is established in the next result.

ORIGINAL PAGE IS
OF POOR QUALITY

Theorem 2.2 Let $y(t;\zeta,u) = (x(t;\zeta,u),x_t(\zeta,u))$, where x is the solution to (2.1) corresponding to $\zeta = (\phi(0),\phi) \in W$ and $u \in H^1((a,b);R^m)$. Then $y(\zeta,u)$ is the unique solution on $[a,b]$ to

$$(2.5) \qquad z(t) = \zeta + \int_a^t \{A(\sigma)z(\sigma) + (B(\sigma)u(\sigma),0)\}d\sigma,$$

for $(B(\sigma)u(\sigma),0) \in Z$. Furthermore, $y(t;\zeta,u) \in Z$ is continuous in $t \in [a,b]$ and uniformly continuous in $\{\zeta,u\} \in W \times H^1((a,b);R^m)$ (in the $Z \times L_2((a,b);R^m)$ topology) uniformly in $t \in [a,b]$.

Remark 2.4 Although we have established an equivalence between an FDE (in the finite-dimensional state variable $x(t)$) and an AEE (in the infinite-dimensional state variable $z(t)$), the infinite dimensionality of the latter is actually inherited from the FDE in that the $L_2(-r,0)$ history of x on $[t-r,t)$ is required before the solution x to the FDE may be determined at t. Thus any computational difficulties associated with (2.5) are not due to the approach we have taken here but, rather, may be traced to the inherent infinite dimensionality of the underlying FDE.

In view of the equivalence established between the FDE and the AEE dynamical systems, we may now restate the control problem P in this abstract setting. That is, we wish to find a control $\bar{u} \in U$ that minimizes

$$\begin{aligned}(2.6) \quad J(u) = {} & (\pi_0 z(b;u) - \kappa(0))Q_0(\pi_0 z(b;u) - \kappa(0)) \\ & + \int_{-r}^{0} (\pi_1 z(b;u) - \kappa)Q_1(\pi_1 z(b;u) - \kappa)(\theta)d\theta \\ & + \int_a^b (\pi_0 z(t;u))Q_2(\pi_0 z(t;u))dt \\ & + \int_a^b \{u(t)Ru(t) + \dot{u}(t)S\dot{u}(t)\}dt,\end{aligned}$$

ORIGINAL PAGE IS
OF POOR QUALITY

where $\pi_0 : Z \to R^n$ and $\pi_1 : Z \to L_2(-r,0)$ are defined by $\pi_0(\xi,\psi) = \xi$ and $\pi_1(\xi,\psi) = \psi$, and $z(t;u)$ satisfies the AEE (2.5). Several properties of P important to the development in §3 are summarized in the following theorem.

Theorem 2.3. Let u_k be given in U, $k = 1,2,\ldots$. Then

(i) the cost functional $J(u_k) \to \infty$ if $|\dot{u}_k| \to \infty$,

(ii) the functional J is weakly lower semicontinuous on $H^1((a,b);R^m)$, and,

(iii) the control problem P has a solution in U.

Proof: Trivial arguments establish (i). For (ii), we will let $\{u_k\} \in U$ be given such that u_k converges weakly to some u in $H^1((a,b);R^m)$, and establish that $J(u) \leq \underline{\lim}_{k\to\infty} J(u_k)$. Rellich's Lemma (the injection of H^1 into L_2 is compact; see, for example, [16]) may be employed to obtain strong convergence of u_k to u in L_2 so that we may utilize continuity properties of z to claim that $z(t;u_k) \to z(t;u)$ in Z, uniformly in t (Theorem 2.2). It immediately follows that

$$(\pi_0 z(b;u_k) - \kappa(0))Q_0(\pi_0 z(b;u_k) - \kappa(0)) \to$$

$$(\pi_0 z(b;u) - \kappa(0))Q_0(\pi_0 z(b;u) - \kappa(0)) \qquad \text{in } R\ ,$$

$$(\pi_1 z(b;u_k) - \kappa)Q_1(\pi_1 z(b;u_k) - \kappa) \to$$

$$(\pi_1 z(b;u) - \kappa)Q_1(\pi_1 z(b;u) - \kappa) \qquad \text{in } L_1((-r,0);R)\ ,$$

and, for $t \in [a,b]$,

$$(\pi_0 z(t;u_k))Q_2(\pi_0 z(t;u_k)) \to (\pi_0 z(t;u))Q_2(\pi_0 z(t;u)) \qquad \text{in } R.$$

ORIGINAL PAGE IS
OF POOR QUALITY

In fact, the last convergence is dominated because we may find $c_0 \in R$ such that

$$|(\pi_0 z(t;u_k))Q_2(\pi_0 z(t;u_k))| \leq c_0^2 |Q_2|$$

due to the fact that $(t,v) \to \pi_0 z(t;v)$ is continuous and hence bounded on the compact set $[a,b] \times \bar{V}$, where $V = \{u_k\}$.

The remaining terms in J are

$$\int_a^b u(t)Ru(t)dt + \int_a^b \dot{u}(t)S\dot{u}(t)dt.$$

We note that for any positive semidefinite $m \times m$ matrix K, the mapping $v \to |v|_K$ defined by

$$|v|_K = \int_a^b v(t)Kv(t)dt$$

is continuous with respect to the L_2 norm and, due to its convexity properties, it is also weakly lower semicontinuous on L_2 (Mazur's Theorem; see [15], p. 422; [2], p. 30). Hence $|u_k|_R \to |u|_R$ since $u_k \to u$ in L_2 and $|\dot{u}|_S \leq \lim_{k\to\infty} |\dot{u}_k|_S$ since $\dot{u}_k \rightharpoonup \dot{u}$ in L_2. The proof of (ii) is complete.

(iii) Since $J(u) \geq 0$, there exists $\alpha \geq 0$ such that $\alpha = \inf_{u\in U} J(u) = \lim_{k\to\infty} J(u_k)$ for some sequence $\{u_k\}$ in U. We must have $|\dot{u}_k| \leq K_0$ for some $K_0 \geq 0$; otherwise,

ORIGINAL PAGE IS
OF POOR QUALITY

there exists a subsequence $\{u_{k_j}\}$ such that $|\dot{u}_{k_j}| \to \infty$ and $\alpha = \lim_{k_j \to \infty} J(u_{k_j}) = \alpha$, contradicting $\alpha \leq J(v) < \infty$ for any $v \in U$.

Finally, making use of the bounds, $|\dot{u}_k| \leq K_0$ and $|u_k(a)| \leq K$, we establish that $\{u_k\}$ is bounded in H^1 so that $\{u_k\}$ must have a subsequence u_{k_j} converging weakly to $\bar{u}$ in $H^1((a.b);R^m)$. In fact, $\bar{u}$ is in U since U closed and convex implies that U is weakly closed in H^1 ([2], p. 30). The observation that $\bar{u}$ is a solution to the control problem P follows immediately from the lower semi-continuity of J since

$$\alpha \leq J(\bar{u}) \leq \lim_{k_j \to \infty} J(u_{k_j}) = \alpha.$$

□

§3 Approximate Optimal Control Problems

Our objective in this section is to construct a sequence of control problems P^N, each of which depends on a finite-dimensional ordinary differential equation (ODE), that will serve to approximate P, a computationally difficult control problem governed by an FDE (an infinite-dimensional state equation). Conventional optimization techniques may then be applied to P^N, applying standard ODE software schemes each time a solution to the state equation is required. Although the approximation framework we develop is based on linear splines, an extension of these ideas to arbitrary order splines is easily accomplished by adding additional smoothness hypotheses on f and making minor modifications in the arguments given below (see the theory developed in [10] on which all of our work here is based).

ORIGINAL PAGE IS
OF POOR QUALITY

For N = 1,2,..., we shall define the linear spline subspace Z^N of Z by

$$Z^N = \{(\psi(0),\psi) \in Z \mid \psi \text{ is a piecewise linear spline with knots at } t_j^N = -j\frac{r}{N},\ j = 0,1,\ldots,N\},$$

and the orthogonal projection (in the $\langle\cdot,\cdot\rangle_\rho$ topology) of Z onto Z^N along $(Z^N)^\perp$ will be denoted by P^N. The nonlinear operator $A^N(t) : Z \to Z^N$ will be defined by $A^N(t) = P^N A(t) P^N$ for $t \in [a,b]$. We remark here that $A^N(t)$ is well-defined since $P^N Z \subseteq W$, W the domain of A(t); in addition, $A^N(t)$ may acutally be viewed as an operator on Z^N because $P^N z^N = z^N$ for any $z^N \in Z^N$. We may now define "approximating" equations on Z^N, N = 1,2,..., by

$$z^N(t) = P^N\zeta + \int_a^t \{A^N(\sigma)z^N(\sigma) + P^N(B(\sigma)u(\sigma),0)\}d\sigma, \quad a \le t \le b, \tag{3.1}$$

each of which may be replaced by an equivalent ODE on Z^N,

$$\begin{cases} \dot{z}^N(t) = A^N(t)z^N(t) + P^N(B(t)u(t),0), & a < t \le b \\ z^N(a) = P^N\zeta, \end{cases} \tag{3.2}$$

Since Z^n is finite-dimensional (each element of Z^N is completely determined by its value at N+1 knots). Corresponding to the new state spaces Z^N and "approximating" equations on Z^N, we may define an "approximate" control problem for each value of N. We shall let P^N denote the problem of finding $\bar{u} \in \mathcal{U}$ that minimizes $J^N(u)$ over all $u \in \mathcal{U}$, where the "approximate" cost functional J^N is defined in a natural way from the form of J in (2.6): Let

ORIGINAL PAGE IS
OF POOR QUALITY

$$J^N(u) = (\pi_0 z^N(b;u) - \kappa^N(0))Q_0(\pi_0 z^N(b;u) - \kappa^N(0))$$

$$+ \int_{-r}^{0} (\pi_1 z^N(b;u) - \kappa^N)\Gamma\,(\pi_1 z^N(b;u) - \kappa^N)(\theta)d\theta$$

$$+ \int_a^b (\pi_0 z^N(t;u))Q_2(\pi_0 z^N(t;u))dt$$

$$+ \int_a^b \{u(t)Ru(t)dt + \dot{u}(t)S\dot{u}(t)\}dt,$$

where $(\kappa^N(0), \kappa^N) = P^N(\kappa(0), \kappa) \in Z^N$ and $t \to z^N(t;u)$ denotes the solution to (3.2) corresponding to $u \in U$.

At this point it is appropriate to characterize properties of $A^N(t)$, P^N, and solutions to (3.2); it is not surprising that most of these properties are inherited from their infinite-dimensional counterparts in §2.1 and §2.2.

Lemma 3.1 Let y and z be given in Z. Then

$$\langle A^N(t)y - A^N(t)z,\ y-z\rangle_\rho \le \omega(t)|y-z|_\rho^2$$

for every N and almost every $t \in [a,b]$. Here ω is defined as in Lemma 2.1 by $\omega(t) = \frac{3}{2}\, m_1(t) + \frac{\nu}{2} + \frac{\nu}{2}\, m_1^2(t)$.

Proof: The result follows immediately from the same dissipative-type estimate on $A(t)$, because

$$\langle A^N(t)y - A^N(t)z,\ y-z\rangle_\rho$$

$$= \langle P^N A(t)P^N y - P^N A(t)P^N z,\ y-z\rangle_\rho$$

$$= \langle A(t)P^N y - A(t)P^N z,\ P^N y - P^N z\rangle_\rho$$

$$\le \omega(t)\ |\ P^N y - P^N z|_\rho^2$$

$$\le \omega(t)\ |\ y-z|_\rho^2$$

ORIGINAL PAGE IS
OF POOR QUALITY

for almost all $t \in [a,b]$.

□

Theorem 3.1 Let $\zeta = (\phi(0),\phi) \in W$ and $u \in U$ be given. Then there exists a unique solution $t \to z^N(t;\zeta,u) \in Z^N$ to (3.2) continuous on $[a,b]$ with the property that the map

$$\{\zeta,u\} \to z^N(t;\zeta,u)$$

is uniformly continuous (in the $Z \times L_2((a,b);R^m)$ topology) uniformly in N and t.

The proof of this result may be found in [8] in the context of a parameter estimation problem and will not be detailed here; the idea is to use a Galerkin-type procedure to rewrite (3.2) as an ODE in spline basis coefficients of $z^N(t)$, so that we obtain existence, uniqueness, and continuous dependence of solutions through an application of conventional ODE theory. For future reference, we briefly describe the linear spline basis elements and the ODE in the basis coefficients, although we refer the reader to [8] for the details of this construction (which is in the spirit of [10]). In the case of the scalar equation (where $Z^N \subseteq R^1 \times L_2((-r,0);R^1)$), linear spline basis elements are given by

$$\hat{e}_j^N = (e_j^N(0),e_j^N), \quad j = 0,1,\ldots,N$$

where e_j^N is the piecewise linear spline in Z^N characterized by $e_j^N(t_i^N) = \delta_{ij}$, δ_{ij} the Kronecker delta symbol. Then any function z^N with $z^N(t) \in Z^N$ can be written $z^N(t) = \sum_{j=0}^{N} w_j^N(t)\hat{e}_j^N$, where $w_j^N(t) \in R^1$; extending to $Z^N \subseteq R^n \times L_2(-r,0)$, we have $z^N(t) = \sum_{j=0}^{N} w_j^N(t)\hat{e}_j^N$ where now $w_j^N(t) \in R^n$. In either case, the ODE in the $n(N+1)$-vector $w^N(t) = \text{col}(w_0^N(t),\ldots,w_N^N(t))$ of basis coefficients becomes

ORIGINAL PAGE IS
OF POOR QUALITY

$$(3.3)\quad \begin{cases} \dot{w}^N(t) = (Q^N)^{-1}\Big(f(t, w_0^N(t), \sum_{j=0}^N w_j^N(t)e_j^N, \sum_{j=0}^N w_j^N(t)e_j^N(-r_1), \\ \qquad\qquad \ldots, \sum_{j=0}^N w_j^N(t)e_j^N(-r_\nu)) + B(t)u(t), 0, \ldots, 0\Big)^T \\ \qquad + (Q^N)^{-1}H_{12}^N w^N(t), \quad a < t \le b \\ w^N(a) = \zeta^N \end{cases}$$

where $\zeta^N = \mathrm{col}(\zeta_0^N, \zeta_1^N, \ldots, \zeta_N^N)$ is defined by $P^N\zeta = \sum_{j=0}^N \zeta_j^N e_j^N$. The $n(N+1)$-square matrices Q^N and H_{12}^N are given by $Q^N = \mathcal{E}^N(0)^T\mathcal{E}^N(0) + \int_{-r}^0 \mathcal{E}^N(\theta)^T\mathcal{E}^N(\theta)\rho(\theta)d\theta$, with $\mathcal{E}^N = (e_0^N, \ldots, e_N^N) \otimes I$, I represents the $n \times n$ identity matrix, $\otimes$ denotes the Kronecker product, and

$$H_{12}^N = \begin{pmatrix} \langle \dot{e}_0^N, e_0^N \rangle & \cdots & \langle \dot{e}_N^N, e_0^N \rangle \\ \vdots & & \vdots \\ \langle \dot{e}_0^N, e_N^N \rangle & \cdots & \langle \dot{e}_N^N, e_N^N \rangle \end{pmatrix} \otimes I,$$

where here $\langle \cdot, \cdot \rangle$ denotes the ρ-weighted $L_2((-r,0);R^1)$ inner product. For $N = 2,3,\ldots$, both Q^N and H_{12}^N have simple three-band matrix forms, exhibited here for the special case of $\nu = 1$ (so $\rho(\theta) \equiv 1$):

ORIGINAL PAGE IS
OF POOR QUALITY

$$Q^N = \frac{r}{N} \begin{pmatrix} \frac{N}{r}+\frac{1}{3} & \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 \\ 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ 0 & 0 & \frac{1}{6} & \frac{1}{3} \end{pmatrix} \otimes I,$$

$$H_{12}^N = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \otimes I.$$

Finally, we need only make minor modifications in the proof of Theorem 2.3 to establish that P^N enjoys many of the same properties as P:

ORIGINAL PAGE IS
OF POOR QUALITY

Theorem 3.2 Let u_k be given in U, $k = 1,2,\ldots$. Then

(i) the N^{th} cost functional $J^N(u_k) \to \infty$ if $|\dot{u}_k| \to \infty$,

(ii) the functional J^N is weakly lower semicontinuous on $H^1((a,b);R^m)$, and,

(iii) P^N has a solution in U.

§3.1 Convergence of State Variable Approximations

From our findings in the preceding sections we are guaranteed the existence of a sequence of optimal pairs $\{(\bar{u}^N, z^N(\bar{u}^N))\}$ where each $\bar{u}^N$ is a solution to P^N and $z^N(\bar{u}^N)$ is the optimal solution to (3.2) corresponding to $\bar{u}^N$. We have yet to establish that the sequence of pairs converges in any meaningful sense to $(\bar{u},z(\bar{u}))$ where $\bar{u}$ is an optimal control for the original control problem P and $z(\bar{u})$ is the corresponding solution to the original AEE (2.5). Fundamental to this endeavor is the requirement that the approximate state variables $z^N(t;u)$ converge to $z(t;u)$ uniformly in $u \in U^1$, where U^1 is a suitably restricted subset of U. To motivate the definition of U^1, we remark that one of our objectives is to demonstrate the convergence of $J^N(\bar{u}^N)$ to $J(\bar{u})$, so that an approximate cost may be computed for N sufficiently large (it is actually $J(\bar{u}^N)$ that is important in practice but it too is approximated by $J^N(\bar{u}^N)$). To this end, any sequence of controls $\{\bar{u}^N\}$; where $\bar{u}^N$ is a solution to P^N, must be characterized by the property that $|\dot{\bar{u}}^N| \leq M$, for some $M \geq 0$, or we will obtain $J^N(\bar{u}^N) \to \infty$ as $|\dot{\bar{u}}^N| \to \infty$. It is not unreasonable then to want to prove state variable convergence uniform in $u \in U^1 \equiv \{u \in U \mid |\dot{u}| \leq M\}$.

In what follows we will establish that solutions $z^N(\zeta,g)$ to

$$(3.4) \qquad z^N(t) = P^N\zeta + \int_a^t \{A^N(\sigma)z^N(\sigma) + P^N(g(\sigma),0)\}d\sigma,$$

ORIGINAL PAGE IS
OF POOR QUALITY

$N = 1,2,\ldots$, converge to $z(\zeta,g)$ the solution to

$$(3.5) \qquad z(t) = \zeta + \int_a^t \{A(\sigma)z(\sigma) + (g(\sigma),0)\}d\sigma$$

uniformly in $g \in \Lambda = \{g \in H^1(a,b) \mid |g(a)| \leq K_1, \ |\dot{g}| < K_2\}$ for all $t \in [a,b]$ and $\zeta \in W$. (There will be no difficulties encountered in replacing $g(t)$ by $B(t)u(t)$ when we return to the control problem formulation.) Our approach will be in the spirit of [8] where convergence is first demonstrated for $\{\zeta,g\}$ in a smooth subset of $W \times \Lambda$.

Define $S = \{(\psi(0),\psi) \in Z \mid \psi \in H^2(-r,0)\}$. For any constant $M \geq 0$ and $\zeta = (\phi(0),\phi)$ in S, let $I(\zeta,M)$ be given by $I(\zeta,M) = \{g \in H^1(a,b) \mid |\dot{g}| \leq M$ and $g(a) = \dot{\phi}(0) - F(a,\phi(0),\phi)\}$. Convergence of $z^N(t;\zeta,g)$ to $z(t;\zeta,g)$ is not difficult to show if ζ is chosen from S and $g \in I(\zeta,M)$, because $z(t;\zeta,g)$ will be sufficiently smooth to allow use of certain standard spline estimates. These remarks are summarized in the two rather technical results that follow.

Lemma 3.2 (i) For any $M \geq 0$, the solution $t \to z(t;\zeta,g)$ to (3.5) corresponding to $\zeta \in S$ and $g \in I(\zeta,M)$ is such that $z(t) \in S$ for all $t \in [a,b]$.

(ii) S is dense in W.

The proof of (i) requires hypotheses (H1) - (H3) and is presented in detail in Lemma 3.2 of [8] (the arguments are applicable here because $\zeta \in S$ and $g \in I(\zeta,M)$ implies that the pair $\{\zeta,g\} \in I$, where I is defined in [8] and differs from $I(\zeta,M)$ defined here). Part (ii) of the lemma is a standard result from the theory of Sobolev spaces.

If $z = (\psi(0),\psi)$ is sufficiently smooth, standard spline results (see, for example [27], [28]) may be invoked to characterize the order of convergence of

ORIGINAL PAGE IS
OF POOR QUALITY

ψ_I^N to ψ and $\mathcal{D}(\psi_I^N)$ to $\mathcal{D}\psi$, where ψ_I^N is the interpolating linear spline function (with knots at t_j^N) for ψ. The convergence rates outlined in the lemma that follows are established in [10], [8] and are derived from these spline estimates and the fact that $|P^N z - z|_\rho \leq |(\psi_I^N(0), {}^N_I) - z|_\rho$ for all N.

Lemma 3.3 Let $z = (\psi(0),\psi)$ be given in S, and denote by $(\psi^N(0),\psi^N)$ the element $P^N z$ of Z^N. Then the following estimates may be obtained for N sufficiently large:

$$(3.6) \qquad |P^N z - z|_\rho \leq \frac{k_1}{N^2} |\mathcal{D}^2\psi|$$

$$(3.7) \qquad |\mathcal{D}\psi^N - \mathcal{D}\psi| \leq \frac{k_2}{N} |\mathcal{D}^2\psi|$$

$$(3.8) \qquad |\psi^N(\theta) - \psi(\theta)| \leq \left(\frac{k_1}{N^2} + \frac{r^{\frac{1}{2}} k_2}{N} \right) |\mathcal{D}^2\psi|, \quad -r \leq \theta \leq 0,$$

where k_1 and k_2 are positive constants independent of N.

We now consider the problem of demonstrating the convergence of $z^N(t;\zeta,g)$ to $z(t;\zeta,g)$ uniformly in $g \in \Lambda$. To this end, we establish preliminary estimates that facilitate the uniform convergence proofs.

Lemma 3.4 Let $M \geq 0$ and $\zeta = (\phi(0),\phi) \in S$ be given, and let $t \to z(t;g) = (x(t;g), x_t(g))$ denote the solution to (3.5) corresponding to ζ and $g \in I(\zeta,M)$, $t \in [a,b]$. Then

(i) the mapping

$$(t,g) \to A(t)z(t;g) = \left(F(t,x(t;g),x_t(g)),\mathcal{D}x_t(g)\right)$$

is continuous from $R \times L_2(a,b)$ to Z, and,

ORIGINAL PAGE IS
OF POOR QUALITY

(ii) there exists a constant $k = k(\zeta,M) > 0$ such that

$$|\mathcal{D}^2 x_t(g)| \leq k$$

for all $g \in I(\zeta,M)$ and almost all $t \in [a,b]$.

Proof: To establish (i), we demonstrate first the uniform continuity of the map

(3.9) $\quad (t,g) \to F(t,x(t;g),x_t(g))$

from $R \times L_2(a,b)$ to Z for $(t,g) \in [a,b] \times \overline{I(\zeta,M)}$ (here $\overline{I(\zeta,M)}$ denotes the closure of $I(\zeta,M)$ in the L_2 topology). Since $I(\zeta,M)$ is bounded in $H^1(a,b)$, (and thus the set is precompact in $L_2(a,b)$), the continuity of z in t and g established in Theorem 2.2 guarantees that

$$\left\{(t,x(t;g),x_t(g),x(t-r_1;g),\ldots,x(t-r_\nu;g)) \;\middle|\; t \in [a,b], g \in \overline{I(\zeta,M)}\right\}$$

is compact in $R \times Z \times R^{n\nu}$. Hypothesis (H2) may be invoked to complete the arguments needed to obtain uniform continuity of the function given in (3.9).

To complete the arguments for (i), let $\varepsilon > 0$ be given and let g, g' be chosen in $I(\zeta,M)$ such that $|g-g'| < \delta$, $\delta = \delta(\varepsilon)$. It follows that

ORIGINAL PAGE IS
OF POOR QUALITY

$$
\begin{aligned}
&\left|\left(F(t,x(t;g),x_t(g)),\mathcal{D}x_t(g)\right) - \left(F(t,x(t;g'),x_t(g')),\mathcal{D}x_t(g')\right)\right|_\rho^2 \\
&\le |F(t,x(t;g),x_t(g)) - F(t,x(t,g'),x_t(g'))|^2 \\
&\quad + \nu \int_{-r}^{0} |\dot{x}(t+\theta;g) - \dot{x}(t+\theta;g')|^2 d\theta \\
&\le |F(t,x(t;g),x_t(g)) - F(t,x(t;g'),x_t(g'))|^2 \\
&\quad + \nu \int_{-r}^{0} |\dot{\phi}(\theta) - \dot{\phi}(\theta)|^2 d\theta \\
&\quad + 2\nu \int_{a}^{b} |F(\sigma,x(\sigma;g),x_\sigma(g)) - F(\sigma,x(\sigma;g'),x_\sigma(g'))|^2 d\sigma \\
&\quad + 2\nu \int_{a}^{b} |g(\sigma) - g'(\sigma)|^2 d\sigma, \\
&< \varepsilon
\end{aligned}
$$

uniformly in t and g, where we have used the continuity of $(t,g) \to F(t,x(t;g),x_t(g))$ just established. These estimates verify the uniform continuity of the map

$$g \to A(t)z(t;g),$$

uniform in t. Finally, the continuity of $A(t)z(t,g)$ in the pair (t,g) follows from an application of the triangle inequality: For (t,g) and (t',g') in $[a,b] \times I(\zeta,M)$,

ORIGINAL PAGE IS
OF POOR QUALITY

$$
\begin{aligned}
&|A(t)z(t;g) - A(t')z(t';g')|_\rho^2 \\
&\le 2|A(t)z(t;g) - A(t')z(t';g)|_\rho^2 \\
&+ 2|A(t')z(t';g) - A(t')z(t';g')|_\rho^2 \\
&\le 2|F(t,x(t;g),x_t(g)) - F(t',x(t';g),x_{t'}(g))|^2 \\
&+ 2\nu \int_{-r}^{0} |\dot{x}(t+\theta;g) - \dot{x}(t'+\theta;g)|^2 d\theta \\
&+ 2|A(t')z(t';g) - A(t')z(t';g')|_\rho^2 ,
\end{aligned}
$$

where the first and third terms may be made arbitrarily small using the previous arguments. In addition, the middle term is small for $|t-t'|$ small since Lemma 3.2(i) may be used to demonstrate that $\dot{x}$ is uniformly continuous in $t \in [a-r,b]$. The desired continuity of

$$(t,g) \to A(t)z(t;g)$$

then obtains.

(ii) Let M be fixed and consider $|\mathcal{D}^2 x_t(g)|$ for all $g \in I(\varsigma,M)$. The differentiability of f (assumed in (H2)) yields

$$
\begin{aligned}
\ddot{x}(t) = {} & f_\sigma(t,x(t),x_t,\ldots) + f_\xi(t,x(t),x_t,\ldots)\dot{x}(t) \\
& + f_\psi[t,x(t),x_t,\ldots;\ \dot{x}_t] \\
& + \sum_{i=1}^{\nu} f_{y_i}(t,x(t),x_t,\ldots)\dot{x}(t-r_i) + \dot{g}(t)
\end{aligned}
$$

for $t \in (a,b)$, where f_δ denotes the Frechet derivative of $f(\sigma,\xi,\psi,y_1,\ldots,y_\nu)$ with respect to δ, $\delta = \sigma,\xi,\psi,y_1,\ldots,y_\nu$. The global Lipschitz property for f

ORIGINAL PAGE IS
OF POOR QUALITY

ensures that these derivatives (excluding f_σ) are bounded, so that, for almost all $t \in (a,b)$,

$$|f_\xi(t,x(t),x_t,\ldots)\dot{x}(t)| \le m_1(t)|\dot{x}(t)|$$

$$|f_\psi[t,x(t),x_t,\ldots;\ \dot{x}_t]| \le m_1(t)|\dot{x}_t|;$$

we obtain similar estimates for f_{y_i}. Thus,

$$|\ddot{x}(t;g)| \le |f_\sigma(t,x(t),x_t,\ldots)| + m_1(t)(1+r^{1/2} + \nu)L + |\dot{g}(t)| \tag{3.10}$$

almost everywhere on $[a,b]$ for $L \equiv \sup\{|\dot{x}(t;g)|;\ t \in [a-r,b],\ g \in \overline{I(\zeta,M)}\}$. That L is finite is easily obtained from the fact that $\dot{x} = \dot{\phi} \in H^1$ on $[a-r,a]$ and, for $t \in (a,b]$,

$$|\dot{x}(t;g)| \le |F(t,x(t,g),x_t(g))| + |g(t)|,$$

where the first term is bounded (uniform in t and g) from arguments developed in the proof of (i), and

$$|g(t)| \le |g(a)| + \int_a^b |\dot{g}(\theta)|d\theta$$

$$\le |\dot{\phi}(0)| + |F(a,\phi(0),\phi)| + (b-a)^{1/2}M.$$

Therefore, from (3.10),

$$|D^2x_t(g)|^2 \le \int_{a-r}^{b} |\ddot{x}(\theta;g)|^2 d\theta$$

$$\le \int_{-r}^{0} |\ddot{\phi}(\theta)|^2 d\theta + \int_a^b [|f_\sigma(\theta,x(\theta;g),\ x_\theta(g),\ldots)| + m_1(\theta)(1+r^{1/2}+\nu)L + |\dot{g}(\theta)|]^2 d\theta$$

$$\le k^2$$

ORIGINAL PAGE IS
OF POOR QUALITY

for all $t \in [a,b]$ and $g \in I(\zeta,M)$, where $k = k(\zeta,M)$ is finite since $|\dot{g}| \leq M$ and (using (H3)) the map $(\theta,\psi) \to f_\sigma(\theta,\psi(0),\psi,\psi(-r_1),\ldots,\psi(-r_\nu))$ is bounded (continuous) over all (θ,ψ) in the compact set $\{(t,x_t(g))|\ t \in [a,b], g \in \overline{I(\zeta,M)}\}$.

□

Lemma 3.5 Let $\zeta = (\phi(0),\phi)$ be given in S, and let z^N and z denote the solutions to (3.4) and (3.5) respectively, corresponding to ζ and $g \in \Lambda$. Then, for any $M \geq 0$,

$$z^N(t;\zeta,g) \to z(t;\zeta,g) \quad \text{as} \quad N \to \infty$$

uniformly in $g \in I(\zeta,M)$ and $t \in [a,b]$.

Proof. Let $\Delta^N(t) = z^N(t) - z(t)$. Then

$$\Delta^N(t) = P^N\zeta - \zeta + \int_a^t \{A^N(\sigma)z^N(\sigma) - A(\sigma)z(\sigma) + P^N(g(\sigma),0) - (g(\sigma),0)\}d\sigma.$$

We apply a commonly used result from analysis: If X is a Hilbert space and if $x : [a,b] \to X$ is given by $x(t) = x(a) + \int_a^t v(\sigma)d\sigma$, then $|x(t)|^2 = |x(a)|^2 + 2\int_a^t \langle x(\sigma),v(\sigma)\rangle d\sigma$ (this is actually a restatement of the well-known result [12, p. 100] that $\frac{d}{dt}\frac{1}{2}|x(t)|^2 = \langle \dot{x}(t),x(t)\rangle$). It follows that

ORIGINAL PAGE IS
OF POOR QUALITY

$$|\Delta^N(t)|_\rho^2 = |(P^N - I)\zeta|_\rho^2 + 2\int_a^t \langle A^N(\sigma)z^N(\sigma) - A^N(\sigma)z(\sigma), \Delta^N(\sigma)\rangle_\rho d\sigma$$

$$+ 2\int_a^t \langle A^N(\sigma)z(\sigma) - A(\sigma)z(\sigma), \Delta^N(\sigma)\rangle_\rho d\sigma$$

$$+ 2\int_a^t \langle (P^N - I)(g(\sigma),0), \Delta^N(\sigma)\rangle_\rho d\sigma$$

$$\leq |(P^N - I)\zeta|_\rho^2 + 2\int_a^t w(\sigma)|\Delta^N(\sigma)|_\rho^2 \, d\sigma$$

$$+ 2\int_a^t |A^N(\sigma)z(\sigma) - A(\sigma)z(\sigma)|_\rho \, |\Delta^N(\sigma)|_\rho d\sigma$$

$$+ 2\int_a^t |(P^N - I)(g(\sigma),0)|_\rho \, |\Delta^N(\sigma)|_\rho d\sigma$$

$$\leq |(P^N - I)\zeta|_\rho^2 + \int_a^b |A^N(\sigma)z(\sigma) - A(\sigma)z(\sigma)|_\rho^2 d\sigma$$

$$+ \int_a^b |(P^N - I)(g(\sigma),0)|_\rho^2 d\sigma + \int_a^b 2(w(\sigma)+1)|\Delta^N(\sigma)|_\rho^2 d\sigma,$$

using $2ab \leq a^2 + b^2$ and the dissipative result for A^N. An application of the Gronwall inequality establishes

$$|\Delta^N(t)|_\rho^2 \leq [\varepsilon_1(N) + \varepsilon_2(N) + \varepsilon_3(N)] \exp \int_a^b 2(w(\sigma)+1)d\sigma,$$

where

$$\varepsilon_1(N) = |(P^N - I)\zeta|_\rho^2 ,$$

$$\varepsilon_2(N) = \int_a^b |A^N(\sigma)z(\sigma) - A(\sigma)z(\sigma)|_\rho^2 d\sigma,$$

and

$$\varepsilon_3(N) = \int_a^b |(P^N - I)(g(\sigma),0)|_\rho^2 d\sigma.$$

ORIGINAL PAGE IS
OF POOR QUALITY

To demonstrate the desired convergence of $z^N(t)$ to $z(t)$ it suffices to show that $\varepsilon_1(N) \to 0$ as $N \to \infty$, uniformly in $g \in I(\zeta,M)$; that the convergence is uniform in t is readily seen.

From Theorem 4.1 of [10] we obtain the strong convergence of P^N to I over Z. It is clear then that $\varepsilon_1(N) \to 0$ and that the integrand in $\varepsilon_3(N)$ also converges to zero uniformly in $g \in \overline{I(\zeta,M)}$. Uniform domination is established by calculating that

$$|(P^N - I)(g(\sigma),0)|_\rho^2 \le 4|(g(\sigma),0)|_\rho^2 = 4|g(\sigma)|^2$$

$$\le 4(|g(a)| + \int_a^\sigma |\dot{g}(\theta)|d\theta)^2$$

$$\le 4(|\dot{\phi}(0)| + |F(a,\phi(0),\phi)| + (b-a)^{1/2}M)^2 ,$$

so that we may claim, using a uniform Dominated Convergence Theorem (see p. 241 of [17]) that, for each $M \ge 0$, $\varepsilon_3(N) \to 0$ as $N \to \infty$, uniformly in $g \in I(\zeta,M)$. Finally, to demonstrate the convergence of $\varepsilon_2(N)$ to zero, we let $(y_\sigma^N(r \quad) \in Z^N$ represent $P^N z(\sigma) = P^N(x_\sigma(0),x_\sigma)$ and observe that

$$|A^N(\sigma)z(\sigma) - A(\sigma)z(\sigma)|_\rho^2$$

$$= |P^N A(\sigma)(y_\sigma^N(0),y_\sigma^N) \quad A(\sigma)(x_\sigma(0),x_\sigma)|_\rho^2$$

$$= |P^N\left(F(\sigma,y_\sigma^N(0) \quad ,\mathcal{D}y_\sigma^N\right) - \left(F(\sigma,x_\sigma(0),x_\sigma),\mathcal{D}x_\sigma\right)|_\rho^2$$

$$\le 2|P^N\left(F(\sigma,y_\sigma^N(0),y_\sigma^N) - F(\sigma,x_\sigma(0),x_\sigma),\mathcal{D}y_\sigma^N - \mathcal{D}x_\sigma\right)|_\rho^2$$

$$+ 2|P^N\left(F(\sigma,x_\sigma(0),x_\sigma),\mathcal{D}x_\sigma\right) - \left(F(\sigma,x_\sigma(0),x_\sigma),\mathcal{D}x_\sigma\right)|_\rho^2$$

$$\le T_1^N(\sigma) + T_2^N(\sigma) + T_3^N(\sigma),$$

ORIGINAL PAGE IS
OF POOR QUALITY

where

$$T_1^N(\sigma) = 2|F(\sigma,y_\sigma^N(0),v_\sigma^N) - F(\sigma,x_\sigma(0),x_\sigma)|^2 ,$$

$$T_2^N(\sigma) = 2\nu|\mathcal{D}y_\sigma^N - \mathcal{D}x_\sigma|^2 ,$$

and

$$T_3^N(\sigma) = 2|P^N\Big(F(\sigma,x_\sigma(0),x_\sigma),\mathcal{D}x_\sigma\Big) - \Big(F(\sigma,x_\sigma(0),x_\sigma),\mathcal{D}x_\sigma\Big)|_\rho^2 .$$

To complete the proof we demonstrate the dominated convergence of $T_i^N(\sigma) \to 0$ uniformly in $g \in I(\zeta,M)$, frequently applying the spline estimates given in Lemma 3.3 (which are valid because $\zeta \in S$ and $g \in I(\zeta,M)$ implies that $z(\sigma) \in S$ for all $\sigma \in [a,b]$). From (H1) and (3.8) we obtain, for N sufficiently large,

$$T_1^N(\sigma) \le 2m_1^2(\sigma)\{|y_\sigma^N(0) - x_\sigma(0)| + |y_\sigma^N - x_\sigma| + \sum_{i=1}^{\nu}|y_\sigma^N(-r_i) - x_\sigma(-r_i)|\}^2$$

$$\le 2m_1^2(\sigma)(1+r^{1/2}+\nu)^2\left(\frac{k_1}{N^2} + \frac{r^{1/2}k_2}{N}\right)^2|\mathcal{D}^2x_\sigma|^2.$$

where $|\mathcal{D}^2x_\sigma|^2 \le k^2$ (a.e. in σ) for all $g \in I(\zeta,M)$ from Lemma 3.4. The spline estimate in (3.7) establishes that, for N sufficiently large,

$$T_2^N(\sigma) \le \frac{2\nu k_2^2}{N^2}|\mathcal{D}^2x_\sigma|^2 \le \frac{2\nu k_2^2k^2}{N^2} ,$$

so that we have, for almost all $\sigma \in [a,b]$, $T_1^N(\sigma) \to 0$ and $T_2^N(\sigma) \to 0$ uniformly in $g \in I(\zeta,M)$; in addition,

$$T_1^N(\sigma) \le 2m_1^2(\sigma)(1 + r^{1/2} + \nu)^2(k_1 + r^{1/2}k_2)^2k^2,$$

$$T_2^N(\sigma) \le 2\nu k_2^2k^2$$

so that each is dominated uniformly in g by an integrable function. Finally,

ORIGINAL PAGE IS
OF POOR QUALITY

$T_3^N(\sigma) \to 0$ as $N \to \infty$ for every $\sigma \in [a,b]$ from the strong convergence of P^N to I on Z. In fact, the convergence is uniform on $V \subseteq Z$, where $V \equiv \{\left(F(\sigma,x(\sigma;g),x_\sigma(g)),\mathcal{D}x_\sigma(g)\right) \mid \sigma \in [a,b], \quad g \in \overline{I(\zeta,M)}\}$ is compact from Lemma 3.4(i). Uniform dominated convergence follows similarly since

$$T_3^N(\sigma) \leq 8 \left|\left(F(\sigma,x(\sigma;g),x_\sigma(g)),\mathcal{D}x_\sigma(g)\right)\right|_\rho^2$$

$$\leq 8 \sup_{v \in V} |v|_\rho^2 < \infty.$$

□

Theorem 3.3 Let $\zeta = (\phi(0),\phi)$ be fixed in W. Then

$$z^N(t;\zeta,g) \to z(t;\zeta,g)$$

as $N \to \infty$ uniformly in $g \in \Lambda$ and $t \in [a,b]$.

Proof: Let $\varepsilon > 0$ be given. From the uniform continuity of the maps

$$(\zeta,g) \to z(t;\zeta,g)$$

$$(\zeta,g) \to z^N(t;\zeta,g),$$

(uniform in t and N), $\delta > 0$ may be determined so that

$$|z(t;\zeta,g) - z(t;\zeta_0,g_0)|_\rho < \varepsilon/3$$

$$|z^N(t;\zeta,g) - z^N(t;\zeta_0,g_0)|_\rho < \varepsilon/3$$

for all $t \in [a,b]$ and $N = 1,2,\ldots$, whenever $|\zeta - \zeta_0| < \delta$, $|g - g_0| < \delta$. Since S is dense in W we may select $\zeta_0 = (\phi_0(0),\phi_0) \in S$ that satisfies this condition.

ORIGINAL PAGE IS
OF POOR QUALITY

In addition, it is a straightforward but tedious exercise to compute M_0 (which depends on ε, ζ_0, and K_1, K_2 from the definition of Λ) such that, for any $g \in \Lambda$, we may construct $g_0 \in I(\zeta_0,M_0)$ that satisfies $|g-g_0| < \delta$. For example, for a given g, one might look for a function g_0 of the form

$$g_0(a) = \dot{\phi}_0(0) - F(a,\phi_0(0),\phi_0),$$

$$g_0(\theta) = \begin{cases} g_0(a) + \left[\dfrac{g(\tau) - g_0(a)}{\tau-a} \right](\theta-a), & a \le \theta \le \tau, \\ g(\theta), & \tau < \theta \le b ; \end{cases}$$

if M_0 is chosen carefully (see pp. 76-78 of [14] for details) then it is not difficult to show that $|\dot{g}_0| \le M_0$ and $|g-g_0| < \delta$ so long as τ is positioned appropriately in $[a,b]$. Therefore it follows from Lemma 3.5 that there exists $N_0 \ge 0$ such that

$$\begin{aligned} &|z^N(t;\zeta,g) - z(t;\zeta,g)|_\rho \\ &\quad \le |z^N(t;\zeta,g) - z^N(t;\zeta_0,g_0)|_\rho \\ &\quad + |z^N(t;\zeta_0,g_0) - z(t;\zeta_0,g_0)|_\rho \\ &\quad + |z(t;\zeta_0,g_0) - z(t;\zeta,g)|_\rho \qquad < \varepsilon \end{aligned}$$

for $N \ge N_0$. The convergence in the middle term is uniform in $g_0 \in I(\zeta_0,M_0)$, so the choice of N_0 depends on $I(\zeta_0,M_0)$, i.e., on ζ (through ζ_0), ε and Λ. It follows immediately that

$$z^N(t;\zeta,g) \to z(t;\zeta,g) \text{ as } N \to \infty$$

uniformly in $g \in \Lambda$ and $t \in [a,b]$.

□

ORIGINAL PAGE IS
OF POOR QUALITY

§3.2 Convergence of Optimal Controls

In §3 we established the existence of an optimal control $\bar{u}^N$ for each approximate control problem P^N, and a solution $z^N(t;\bar{u}^N) = (x^N(t;\bar{u}^N), x_t^N(\bar{u}^N))$ to (3.2) associated with $\bar{u}^N$. Our final result indicates when the approximate problem P^N may be used to compute numerical solutions for the original control problem P.

Theorem 3.4 Let $\{\bar{u}^N\}$ denote a sequence of controls in U where $\bar{u}^N$ is optimal for the approximating control problem P^N. Then there is a control $\bar{u} \in U$ and a subsequence $\{\bar{u}^{N_k}\}$ of $\{\bar{u}^N\}$ such that

(i) $\bar{u}^{N_k} \to \bar{u}$ in $H^1((a,b);R^m)$,

(ii) $\bar{u}^{N_k} \to \bar{u}$ in $L_2((a,b);R^m)$,

(iii) $x^{N_k}(t;\bar{u}^{N_k}) \to x(t;\bar{u})$ uniformly in $t \in [a,b]$, where $x^{N_k}(t) = \Pi_0 z^{N_k}(t)$ and $x(t) = \Pi_0 z(t)$ are the first components of the solutions to (3.2) and (2.5), respectively,

(iv) $J^{N_k}(\bar{u}^{N_k}) \to J(\bar{u})$, and

(v) $\bar{u}$ is a solution to the original control problem P.

Proof: We first remark that, for any $v \in U$, $J^N(v) \to J(v)$; this is trivially true because $P^N(\kappa(0),\kappa) \to (\kappa(0),\kappa)$ in Z and $z^N(t;v) \to z(t,v)$ in Z, uniformly in $t \in [a,b]$, from Theorem 3.3.

The sequence $\{\bar{u}^N\}$ must have the property that $|\dot{\bar{u}}^N| \leq M$, for some $M \geq 0$, or there would exist a subsequence $\{\bar{u}^{N_k}\}$ such that $|\dot{\bar{u}}^{N_k}| \to \infty$, and thus

ORIGINAL PAGE IS
OF POOR QUALITY

$J^{N_k}(\bar{u}^{N_k}) \to \infty$, which contradicts the inequality

$$J^{N_k}(\bar{u}^{N_k}) \leq J^{N_k}(v) \to J(v) < \infty$$

for any $v \in U$. This result and the definition of U establish that $\{\bar{u}^N\}$ is bounded in $H^1((a,b);R^m)$. We may therefore extract a weakly convergent subsequence $\{\bar{u}^{N_k}\}$ such that $\bar{u}^{N_k} \rightharpoonup \bar{u}$ in H^1, for some $\bar{u} \in U$, proving (i). The proof of (ii) is an application of Rellich's Lemma. To demonstrate the convergence needed in (iii), we employ Theorems 2.2 and 3.3 to observe that

$$|x^{N_k}(t;\bar{u}^{N_k}) - x(t;\bar{u})| \leq |z^{N_k}(t;\bar{u}^{N_k}) - z(t;\bar{u})|_\rho$$

$$\leq |z^{N_k}(t;\bar{u}^{N_k}) - z(t;\bar{u}^{N_k})|_\rho + |z(t;\bar{u}^{N_k}) - z(t;\bar{u})|_\rho \to 0$$

as $N_k \to \infty$ uniformly in $t \in [a,b]$, where convergence in the first term is uniform in $\bar{u}^{N_k}$ since $B\bar{u}^{N_k} \in \{g=Bv \mid v \in U, |\dot{v}| \leq M\} \subseteq \Lambda$.

Finally, the weak lower semicontinuity of J and the convergence of $z^{N_k}(t;\bar{u}^{N_k})$ to $z(t;\bar{u}^{N_k})$ uniform in $\{\bar{u}^{N_k}\}$ may be used to demonstrate

$$\begin{aligned} J(\bar{u}) &\leq \underline{\lim}_{N_k \to \infty} J(\bar{u}^{N_k}) \\ &= \underline{\lim}_{N_k \to \infty} J^{N_k}(\bar{u}^{N_k}) \\ &\leq \overline{\lim}_{N_k \to \infty} J^{N_k}(\bar{u}^{N_k}) \\ &\leq \overline{\lim}_{N_k \to \infty} J^{N_k}(v) \\ &= J(v) \end{aligned}$$

ORIGINAL PAGE IS
OF POOR QUALITY

for any $v \in U$. Thus, the control $\bar{u} \in U$ must be optimal for the original control problem P. Repeating the same arguments with $\bar{u}$ in place of v we also find

$$J(\bar{u}) = \lim_{N_k \to \infty} J^{N_k}(\bar{u}^{N_k}) .$$

□

§4 Numerical Results

In the final section we present our numerical findings for a number of inertial control problems governed by functional differential equations. To illustrate the method developed in previous sections we chose examples that were representative of problems found frequently in applications or that were of special interest because of their difficult or commonly occuring nonlinear structure. In Example 4.1 we consider the problem of controlling the Mach number in the test chamber of the liquid nitrogen wind tunnel currently being built at NASA Langley Research Center. Problems with nonlinearities of the form $\sin(x)$ and $x \sin(x)$ are examined in Examples 4.2 and 4.3, where it is interesting to note that performance of the spline approximations for each is excellent in spite of the fact that the second nonlinearity is only locally Lipschitz (and thus the FDE in question does not satisfy the hypotheses detailed above). In the control problem presented in Example 4.4, we consider a delay equation with a nonlinearity of the Michaelis - Menten type; such an equation is commonly used to approximate velocity of reaction in models involving enzyme-mediated chemical reactions. The example also serves to illustrate convergence in models with point delays in the nonlinear part of the FDE, which would be important if one desired to model the transport (with a time lag r) of substrate/product molecules in and out of a compartment where enzymatic reaction takes place. Finally, in

ORIGINAL PAGE IS
OF POOR QUALITY

Example 4.5 we compare the spline-based approximation scheme to the averaging approximation method developed in [3], using an example from that reference. Although the control problem is not of inertial type, we are able to successfully apply our method and demonstrate that the spline approximations exhibit better performance for this example.

In the examples that follow we consider the open loop control problem of finding $\bar{u} \in U = \{u \in H^1((0,T);R) \mid u(0) = 0\}$ that minimizes

$$\tilde{J}(u) = \frac{1}{2}x(T)\tilde{Q}_0x(T) + \frac{1}{2}\int_0^T \{x(t)\tilde{Q}_1x(t) + u(t)\tilde{R}u(t) + (\dot{u}(t))^2\}dt \tag{4.1}$$

where x satisfies

$$\begin{cases} \dot{x}(t) = \tilde{A}_0x(t) + \tilde{A}_1x(t-r) + \tilde{F}(t,x(t),x(t-r)) + \tilde{B}u(t), \ 0 \le t \le T, \\ x_0 = \phi \in H^1(-r,0) , \end{cases} \tag{4.2}$$

and $\tilde{A}_0$, $\tilde{A}_1$, $\tilde{Q}_0$, and $\tilde{Q}_1$ are $n \times n$ matrices, $\tilde{R} \in R$, and $\tilde{B}$ is an $n \times 1$ matrix. To simplify our calculations we replace (4.2) by an equivalent delay equation in the $(n+1)$-vector $y = (x,u)^T$,

$$\begin{cases} \dot{y}(t) = A_0y(t) + A_1y(t-r) + F(t,y(t), y(t-r)) + Bv(t), \ 0 \le t \le T, \\ y_0 = \begin{pmatrix} \phi \\ 0 \end{pmatrix} \end{cases} \tag{4.3}$$

where $v(t) = \dot{u}(t)$ is now treated as the control function; here the $(n+1)$-square matrices A_0 and A_1 are given by

$$A_0 = \begin{pmatrix} \tilde{A}_0 & \tilde{B} \\ \\ \Theta^T & 0 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} \tilde{A}_1 & \Theta \\ \Theta^T & 0 \end{pmatrix},$$

ORIGINAL PAGE IS
OF POOR QUALITY

and

$$B = \begin{pmatrix} \Theta \\ 1 \end{pmatrix}$$

where Θ is the zero vector in R^n. Similarly, the nonlinearity is defined to be

$$F(t,y(t),y(t-r)) = \begin{pmatrix} \tilde{F}(t,x(t),x(t-r)) \\ 0 \end{pmatrix}.$$

The cost functional associated with (4.3) may now be written as

$$(4.4) \qquad J(v) = \frac{1}{2} y(T)Q_0 y(T) + \frac{1}{2}\int_0^T \{y(t)Q_1 y(t) + v^2(t)\}dt$$

where

$$Q_0 = \begin{pmatrix} \tilde{Q}_0 & \Theta \\ \Theta^T & 0 \end{pmatrix}, \quad Q_1 = \begin{pmatrix} \tilde{Q}_1 & \Theta \\ \Theta^T & \tilde{R} \end{pmatrix}$$

are (n+1)-square matrices. We turn now to the problem of determining a numerical solution for P^N, the N^{th} approximating control problem for (4.3), (4.4). From (3.3) and the discussion following Theorem 3.1, the approximate control problem is that of finding $\bar{v}^N$ that minimizes

$$(4.5) \qquad J(v) = \frac{1}{2} w_0^N(T)Q_0 w_0^N(T) + \frac{1}{2}\int_0^T \{w_0^N(t)Q_1 w_0^N(t) + v^2(t)\}dt$$

ORIGINAL PAGE IS
OF POOR QUALITY

over all $v \in L_2((0,T);R)$; here w_0^N (the Z^N basis coefficients for $y = (x,u)^T$) is the first $(n+1)$-vector component of $w^N = (w_0^N,\dots,w_N^N)^T \in R^{(n+1)(N+1)}$ satisfying

$$(4.6)\quad \begin{cases} \dot{w}^N(t) = (Q^N)^{-1}\begin{pmatrix} A_0 w_0^N(t) + A_1 w_N^N(t) + F(t,w_0^N(t),w_N^N(t)) + Bv(t) \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (Q^N)^{-1} H_{12}^N w^N(t), \quad 0 \le t \le T \\ w^N(0) = \zeta^N, \end{cases}$$

where ζ^N is defined from the Z^N basis coefficients of $P^N(\phi(0),\phi)$, and Q^N and H_{12}^N were given in §3.

To obtain an optimal control $\bar{v}^N$ (and the associated optimal trajectory $\bar{w}_0^N$) for P^N, we applied standard computational routines (such as gradient and conjugate gradient schemes; see [23]) to necessary conditions for quadratic control problems governed by nonlinear ODE's (see [22], for example). All computations were executed on the CDC 6600 computer at Southern Methodist University, using a software package developed by Dr. D. Reber in 1977 while he was a student at Brown University. A full description of the optimization package and convergence criteria may be found in [3; Section 4]; integration of all ODE's was accomplished using Gill's modification of a standard fourth-order Runge-Kutta method.

Since analytic solutions for these nonlinear problems are not available, an independent method of checking the validity of our results is desirable. To this end we worked directly with the FDE-governed control problem given by (4.3) and (4.4), and applied necessary conditions for optimality of delay system problems (detailed in Theorem VII.2.31 of [24]). If $(\bar{y},\bar{v})$ is an extremal pair for (4.3), (4.4), the necessary conditions ensure that a multiplier $\lambda(t) \in R^{n+1}$ exists such that $\bar{v}$ maximizes

ORIGINAL PAGE IS
OF POOR QUALITY

$$(4.7)\qquad H = -\frac{1}{2}v^2(t) + \lambda(t)\{A_0\bar{y}(t) + A_1\bar{y}(t-r) + F(t,\bar{y}(t), \bar{y}(t-r)) + Bv(t)\}$$

(i.e., $\bar{v} = B\lambda$) and λ satisfies

$$(4.8)\qquad \begin{cases} \dot{\lambda}(t) = -(A_0 + \mathcal{D}_2F(t,\bar{y}(t),\bar{y}(t-r)))^T\lambda(t) \\ \qquad\qquad - (A_1 + \mathcal{D}_3F(t+r,\bar{y}(t+r),\bar{y}(t)))^T\lambda(t+r) \\ \qquad\qquad + Q_1\bar{y}(t), \quad 0 \le t \le T, \\ \lambda(T) = -Q_0\bar{y}(T), \\ \lambda(t) = 0, \quad t > T, \end{cases}$$

where $\mathcal{D}_iF$ denotes the derivative of F with respect to its i^{th} component, i = 1,2,3. Therefore, to solve the original control problem directly, we must solve a mixed advanced/delayed two-point boundary value problem (TPBVP) that consists of equations (4.3) and (4.8). For our purposes this was accomplished using one of the fourth-order block methods developed by Tavernini [29] for solving FDE's. We were able to construct an approximate numerical solution $(\bar{y},\bar{v})$ to the original control problem by first integrating (4.8) backwards to obtain λ (using the spline approximation $\bar{w}_0^{32}$ in place of $\bar{y}$ in (4.8)); we then used the method in [29] once again to integrate (4.3) forward in time to obtain $\bar{y}$ (substituting $\bar{v} = B\lambda$). One need then only check to see if $\lambda(T) \approx -Q_0\bar{y}(T)$ to determine whether the pair $(\bar{y},\bar{v})$ provides an adequate standard by which we may compare the spline approximations $(\bar{w}_0^N, \bar{v}^N)$. As will be seen in Examples 4.2 through 4.5, these values provide a rough, but independent, check on our approximate numerical solutions $(\bar{w}_0^N, \bar{v}^N)$.

ORIGINAL PAGE IS
OF POOR QUALITY

Example 4.1 Our first example is an open loop control problem motivated by the need to regulate the Mach number in the test chamber of the liquid nitrogen wind tunnel (National Transonic Facility - NTF) under construction at NASA Langley Research Center in Virginia. The NTF will be able to generate an air-stream that can reach a maximum of Mach 1.2 in speed, a feat accomplished through the use of high speed fans and nitrogen gas maintained at very low temperatures and high pressure. The liquid nitrogen that is sprayed into the tunnel upstream of the fan section acts primarily to regulate temperature, while pressure control is facilitated by venting nitrogen gas to the outside of the tunnel; Mach number is primarily regulated by fan motor speed, although subtle variations in Mach number are more easily controlled by making changes in inlet guide vanes in the fan section. These features of the wind tunnel are depicted in Figure 1.

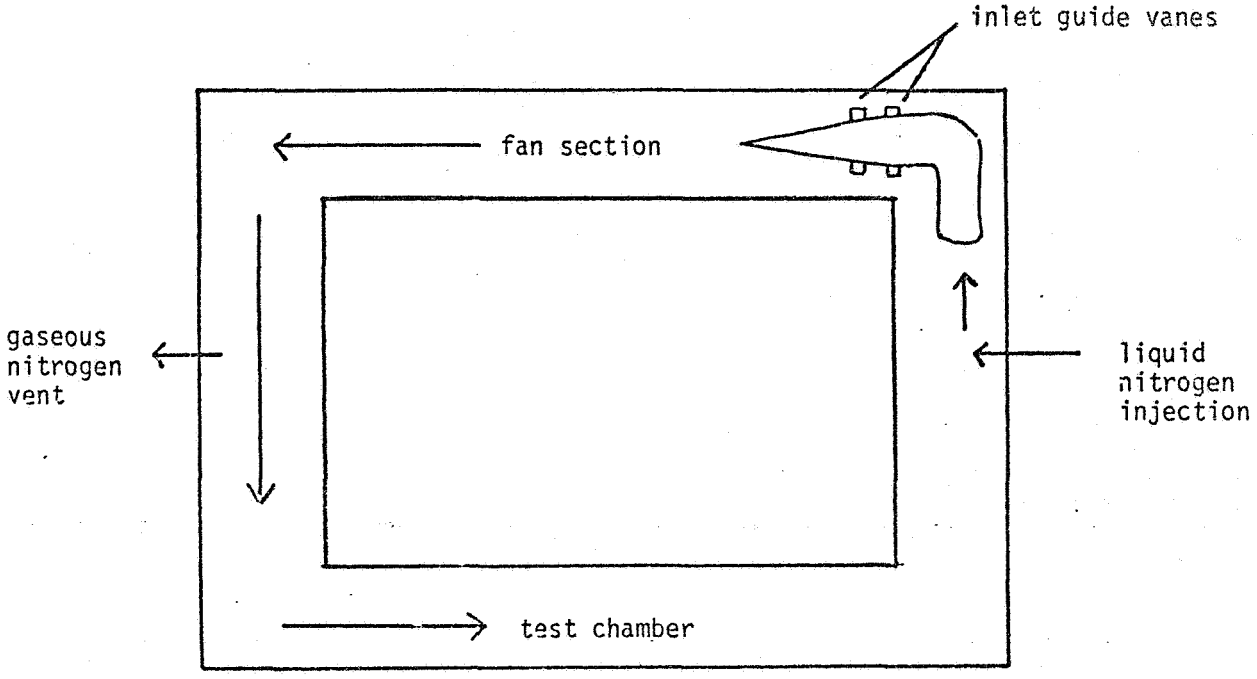


Figure 1

ORIGINAL PAGE IS
OF POOR QUALITY

To control the Mach number, we first need an adequate model for fluid flow in the tunnel and test chamber. Unfortunately, standard Navier-Stokes partial differential equations of fluid flow are extremely difficult to solve computationally and thus would be of limited value when used in conjunction with the design of an optimal controller; a better alternative appears to be a lumped-parameter delay differential equation model [1] in which the delays represent transport lags of the fluid as it moves from one section of the tunnel to the other. In [1], a model of the relationship between Mach number and guide vane angle (GVA) is given by

$$\tau\dot{M}(t) + M(t) = k\theta(t-r) \tag{4.9}$$

where $M(t)$ represents the perturbation of the Mach number from a given set point and $\theta(t)$ is the perturbation of the GVA from a steady state GVA. For this example, we will consider the regulation of Mach number from .8 to a set point of .9, so that the control problem consists of steering $M(t)$ from an initial value of -.1 to the terminal value 0.0 in an efficient manner. The time delay r is approximately .3 sec and, with a Mach number set point of .9, the constants k and τ are given by $-0.0117\ \text{deg}^{-1}$ and 1.964 sec respectively. Change in GVA is initiated by a guide vane actuator; the dynamics of this process are approximated in [1] by the following ODE in θ and θ_A, where θ_A is the perturbation of the guide vane actuator from a steady state value:

$$\ddot{\theta}(t) + 2\zeta\omega\dot{\theta}(t) + \omega^2\theta(t) = \omega^2\theta_A(t)\ . \tag{4.10}$$

The constants ζ and ω are given by .8 and 6.0 rad/sec., respectively. For theoretical reasons discussed in [1], it is desirable to let the guide vane

ORIGINAL PAGE IS
OF POOR QUALITY

actuator rate ($\dot{\theta}_A$) be the control variable v(t). Combining equations (4.9) and (4.10) we obtain a first order linear delay differential equation in $x(t) = (M(t),\theta(t),\dot{\theta}(t),\theta_A(t))^T$, given by

$$\dot{x}(t) = A_0x(t) + A_1x(t-r) + Bv(t) \tag{4.11}$$

where

$$A_0 = \begin{bmatrix} -\frac{1}{\tau} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -\omega^2 & -2\zeta\omega & \omega^2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 0 & \frac{k}{\tau} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$B = (0,0,0,1)^T.$$

Reasonable initial conditions for x are given by

$$x_0 = (-0.1,\ 8.547,\ 0.0,\ 8.547),$$

and a steady state value of 1.93 is given for both the GVA and the guide vane actuator. For the Mach number control problem, the cost functional is chosen

ORIGINAL PAGE IS
OF POOR QUALITY

to be

$$(4.12) \qquad J(v) = \frac{1}{2} M^2(12) + \frac{1}{2} \int_0^{12} \{10^4 M^2(t) + v^2(t)\}dt$$

where $v = \dot{\theta}_A \in L_2(0,12;R)$.

Although time-consuming (CPU time = 252 sec), the spline approximation for N = 6 gives excellent results when compared to the findings in [1], where the approach was to discretize (4.11) by finite-difference techniques. Graphs of the spline-based approximation for the optimal trajectories of Mach number and GVA may be found in Figures 2 and 3, where our findings are contrasted with those of [1].

Example 4.2 Consider the problem of finding $\bar{u} \in U$ that minimizes the cost functional

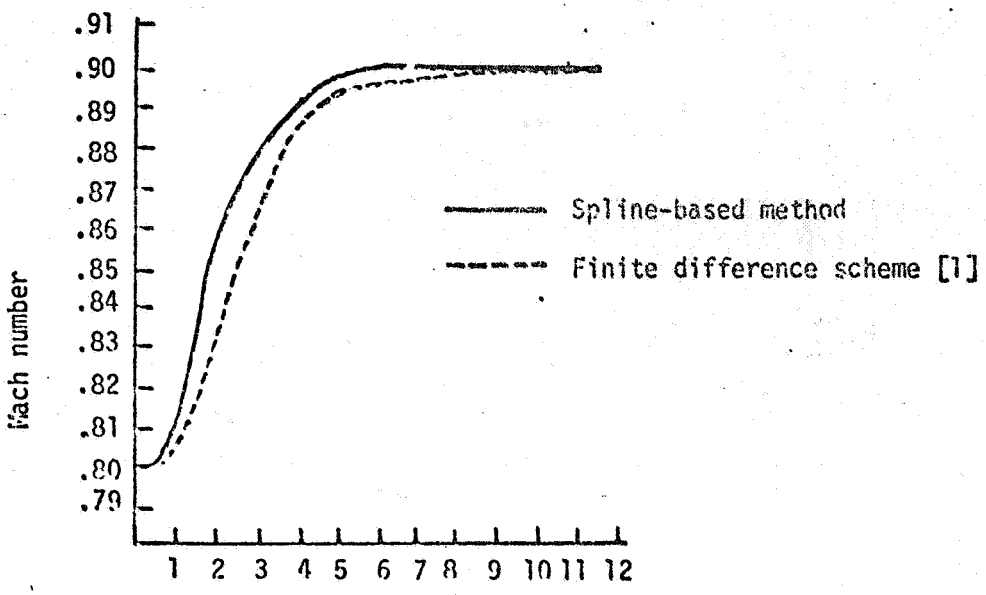
$$J(u) = \frac{1}{2} x^2(2) + \frac{1}{2} \int_0^2 \dot{u}^2(t)dt$$

where $\bar{u}$ and $\bar{x} = \bar{x}(\bar{u})$ satisfy

$$\begin{cases} \dot{x}(t) = \frac{1}{2} t^2 \sin x(t) + x(t-1) + u(t), & 0 \le t \le 2, \\ \phi(\theta) = 1, & -1 \le \theta \le 0. \end{cases}$$

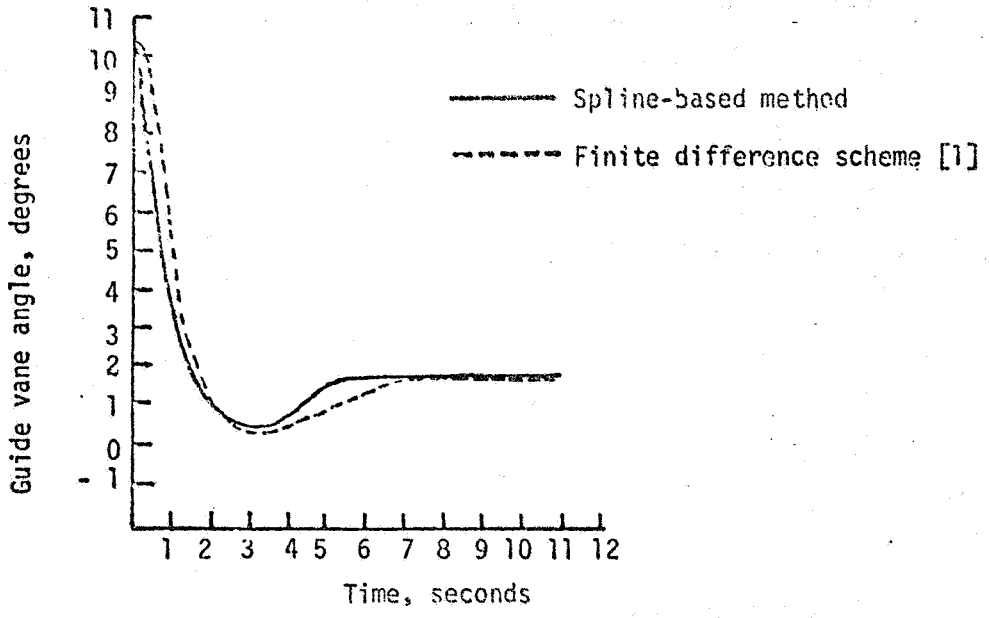
The results of our computations are summarized in Table 1, where we present approximate cost $\bar{J}^N$, the spline approximations $\bar{x}^N$ (for the optimal trajectory) and $\bar{u}^N$ (for the optimal control), where N is taken to be 4, 8, 20 and 32. "True" extremals $\bar{x}$ and $\bar{u}$ were generated from the TPBVP described in the introduction to this section, with

ORIGINAL PAGE IS
OF POOR QUALITY



Example 4.1: Mach Number

FIGURE 1



Example 4.1: GVA

FIGURE 2

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 1

Results for Example 4.2

Time	$\bar{u}^4$	$\bar{u}^8$	$\bar{u}^{20}$	$\bar{u}^{32}$	$\bar{u}$
0.00	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	-0.6674	-0.6899	-0.6873	-0.6874	-0.6858
0.50	-1.2200	-1.2379	-1.2327	-1.2325	-1.2291
0.75	-1.6463	-1.6611	-1.6551	-1.6550	-1.6494
1.00	-1.9674	-1.9758	-1.9726	-1.9725	-1.9645
1.25	-2.2056	-2.1979	-2.1996	-2.1997	-2.1891
1.50	-2.3681	-2.3385	-2.3462	-2.3462	-2.3332
1.75	-2.4558	-2.4072	-2.4226	-2.4232	-2.4087
2.00	-2.4786	-2.4239	-2.4445	-2.4453	-2.4303

Time	$\bar{x}^4$	$\bar{x}^8$	$\bar{x}^{20}$	$\bar{x}^{32}$	$\bar{x}$
0.00	1.0000	1.0000	1.0000	1.0000	1.0000
0.25	1.1634	1.1631	1.1633	1.1633	1.1636
0.50	1.1963	1.1866	1.1875	1.1875	1.1889
0.75	1.1314	1.1174	1.1195	1.1194	1.1228
1.00	1.0168	0.9957	0.9981	0.9979	1.0041
1.25	0.8969	0.8717	0.8758	0.8756	0.8862
1.50	0.7926	0.7727	0.7756	0.7754	0.7927
1.75	0.7031	0.6855	0.6880	0.6878	0.7162
2.00	0.6223	0.6097	0.6088	0.6085	0.6564

$\bar{J}^N$	2.4739	2.4605	2.4567	2.4560	

ORIGINAL PAGE IS
OF POOR QUALITY

$$\lambda(2) = \begin{pmatrix} -0.6080 \\ 0.0 \end{pmatrix}$$

$$-Q_0\bar{y}(2) = \begin{pmatrix} -0.6564 \\ 0.0 \end{pmatrix} ;$$

as in the examples to follow, we see that the values of $\bar{x}$ and $\bar{u}$ calculated in this way are rough approximations to true extremals.

Example 4.3 In this example, we handle a system containing a nonlinearity satisfying only a local Lipschitz condition. The problem is to minimize

$$J(u) = \frac{1}{2}x^2(2) + \frac{1}{2}\int_0^2 \{x^2(t) + u^2(t) + \dot{u}^2(t)\}dt$$

over $u \in U$ subject to

$$\begin{cases} \dot{x}(t) = x(t-1) + x(t)\sin x(t) + u(t) & 0 \le t \le 2 \\ \phi(\theta) = 10, & -1 \le \theta \le 0. \end{cases}$$

The spline approximations (x^N, u^N), $N = 2,4,8,16,32$, for $(\bar{x},\bar{u})$ are given in Table 2, with "optimal" values of $(\bar{x},\bar{u})$ determined from the TPBVP; in this example,

$$\lambda(2) = \begin{pmatrix} -10.3444 \\ 0.0 \end{pmatrix}$$

$$-Q_0\bar{y}(2) = \begin{pmatrix} -10.3457 \\ 0.0 \end{pmatrix} .$$

TABLE 2

ORIGINAL PAGE IS
OF POOR QUALITY

Results for Example 4.3

Time	$\bar{u}^2$	$\bar{u}^4$	$\bar{u}^8$	$\bar{u}^{16}$	$\bar{u}^{32}$	$\bar{u}$
0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	-0.4320	-0.4693	-0.4794	-0.4842	-0.4882	-0.4920
0.50	-0.8120	-0.8646	-0.8714	-0.8776	-0.8832	-0.8875
0.75	-1.1401	-1.1767	-1.1875	-1.1943	-1.2003	-1.2039
1.00	-1.4280	-1.4319	-1.4421	-1.4490	-1.4549	-1.4550
1.25	-1.6901	-1.6699	-1.6755	-1.6808	-1.6861	-1.6798
1.50	-1.9371	-1.9166	-1.9118	-1.9136	-1.9177	-1.9064
1.75	-2.1616	-2.1571	-2.1431	-2.1409	-2.1452	-2.1334
2.00	-2.2910	-2.2932	-2.2807	-2.2785	-2.2840	-2.2757

Time	$\bar{x}^2$	$\bar{x}^4$	$\bar{x}^8$	$\bar{x}^{16}$	$\bar{x}^{32}$	$\bar{x}$
0.00	10.0000	10.0000	10.0000	10.0000	10.0000	10.0000
0.25	10.4570	10.4639	10.4631	10.4609	10.4605	10.4609
0.50	10.5118	10.5018	10.4959	10.4949	10.4941	10.4938
0.75	10.4837	10.4585	10.4589	10.4579	10.4569	10.4564
1.00	10.4490	10.4279	10.4204	10.4167	10.4149	10.4140
1.25	10.4201	10.4204	10.4203	10.4224	10.4210	10.4215
1.50	10.3933	10.4095	10.4135	10.4114	10.4103	10.4116
1.75	10.3658	10.3787	10.3766	10.3784	10.3776	10.3793
2.00	10.3421	10.3428	10.3457	10.3452	10.3444	10.3457

$\bar{J}^N$	165.707	165.700	165.702	165.694	165.692	

ORIGINAL PAGE IS
OF POOR QUALITY

Example 4.4 We now consider a nonautonomous problem with a Michaelis-Menten type nonlinearity involving the point delay $x(t-r)$. We wish to find $\bar{u} \in U$ that minimizes

$$J(u) = \frac{1}{2} x^2(2) + \frac{1}{2} \int_0^2 \dot{u}^2(t)dt$$

where $\bar{u}$ and $\bar{x}$ satisfy

$$\begin{cases} \dot{x}(t) = - tx(t) + \dfrac{2}{1+x(t-1)} + u(t), & 0 \leq t \leq 2, \\ \phi(\theta) = 10, & -1 \leq \theta \leq 0 . \end{cases}$$

In this case, the numerical solution of the TPBVP yields $\bar{x}$ and $\bar{u}$ such that

$$\lambda(2) = \begin{pmatrix} - 1.0444 \\ 0.0 \end{pmatrix}$$

$$- Q_0 \bar{y}(2) = \begin{pmatrix} - 1.0166 \\ 0.0 \end{pmatrix} ;$$

the values of $(\bar{x},\bar{u})$ as well as the spline approximations $(\bar{x}^N , \bar{u}^N)$ for $N = 2,4,8,16$ and 32 are presented in Table 3.

Example 4.5 As a final problem we take Example 4.4 from [3] and compare the averaging approximation method detailed there to the spline approximation scheme outlined in this paper. Although the control problem in question is not of inertial type (i.e., neither U nor the form of J requires that $\dot{u}$ be bounded), we are still able to apply our method with success. The problem is to find $\bar{u} \in U = L_2((0,2);R)$ that minimizes

TABLE 3

ORIGINAL PAGE IS
OF POOR QUALITY

Results for Example 4.4

Time	$\bar{u}^2$	$\bar{u}^4$	$\bar{u}^8$	$\bar{u}^{16}$	$\bar{u}^{32}$	$\bar{u}$
0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	-0.1429	-0.1530	-0.1576	-0.1598	-0.1611	-0.1620
0.50	-0.2887	-0.3051	-0.3091	-0.3118	-0.3137	-0.3153
0.75	-0.4309	-0.4478	-0.4523	-0.4549	-0.4567	-0.4590
1.00	-0.5629	-0.5788	-0.5847	-0.5870	-0.5882	-0.5911
1.25	-0.6782	-0.6966	-0.7027	-0.7044	-0.7050	-0.7084
1.50	-0.7704	-0.7953	-0.8013	-0.8024	-0.8024	-0.8060
1.75	-0.8324	-0.8641	-0.8713	-0.8722	-0.8720	-0.8758
2.00	-0.8573	-0.8901	-0.8987	-0.9001	-0.8999	-0.9037

Time	$\bar{x}^2$	$\bar{x}^4$	$\bar{x}^8$	$\bar{x}^{16}$	$\bar{x}^{32}$	$\bar{x}$
0.00	10.0000	10.0000	10.0000	10.0000	10.0000	10.0000
0.25	9.7467	9.7291	9.7207	9.7178	9.7171	9.6915
0.50	8.9221	8.8599	8.8405	8.8356	8.8340	8.7874
0.75	7.6345	7.5383	7.5170	7.5108	7.5088	7.4489
1.00	6.0770	5.9843	5.9646	5.9587	5.9566	5.8925
1.25	4.4723	4.4159	4.4012	4.3967	4.3950	4.3345
1.50	3.0157	3.0050	3.0001	2.9984	2.9976	2.9464
1.75	1.8363	1.8612	1.8671	1.8690	1.8692	1.8297
2.00	0.9837	1.0277	1.0407	1.0437	1.0444	1.0166

$\bar{J}^N$	0.6956	0.7548	0.7726	0.7767	0.7777	

ORIGINAL PAGE IS
OF POOR QUALITY

$$J(u) = \frac{1}{2}\, x^2(2) + \frac{1}{2}\int_0^2 \{x^2(t) + u^2(t)\}dt$$

subject to

$$\begin{cases} \dot{x}(t) = x(t)\sin x(t) + x(t-1) + u(t), & 0 \le t \le 2, \\ \phi(\theta) = \begin{cases} 10(\theta+1), & -1 \le \theta \le -0.5, \\ -10\theta\ , & -0.5 \le \theta \le 0. \end{cases} \end{cases}$$

From Tables 5 and 6 of [3] it is evident that the averaging scheme works quite well, but, as we shall demonstrate, the spline-based method exhibits even better performance for this example. At this point we wish to remark that the "true" extremals $(\bar{x},\bar{u})$ used for comparison in [3] were determined from the TPBVP, where an incorrect version of Tavernini's block method was used to integrate the delayed-advanced equations (in fact, the error is a misprint on p. 793 of Tavernini's original paper [29] where, in the fourth line of the calculations for $x_3(s)$, $2F_h$ and $4F_h$ should be separated by a minus sign). In this particular example the errors involved are slight; they are more substantial for Example 4.2 of the same reference.

In Table 4, the averaging approximations $\bar{x}^N_{AVE}$ and $\bar{u}^N_{AVE}$ are given for N = 8 and 32. Similarly, we provide results for the spline-based method by summarizing our findings for $\bar{x}^N_{SPL}$, $\bar{u}^N_{SPL}$, N = 4, 8, and 32. Both are compared against "optimal" solutions $\bar{x}$ and $\bar{u}$ (using the correct version of Tavernini's method), where, from the TPBVP, it is determined that

$$\lambda(2) = \bar{u}(2) = -\ 0.3230$$

and

$$-\ \bar{x}(2) = -\ 0.3053.$$

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 4

Results for Example 4.5

Time	$\bar{u}^8_{AVE}$	$\bar{u}^{32}_{AVE}$	$\bar{u}^4_{SPL}$	$\bar{u}^8_{SPL}$	$\bar{u}^{32}_{SPL}$	$\bar{u}$
0.00	-2.3083	-2.3032	-2.3331	-2.3024	-2.2967	-2.3028
0.25	-2.1804	-2.2706	-2.3707	-2.3230	-2.3101	-2.3164
0.50	-1.9855	-2.1972	-2.2820	-2.3034	-2.3126	-2.3189
0.75	-1.5751	-1.6908	-1.6921	-1.7294	-1.7424	-1.7470
1.00	-1.1446	-1.1318	-1.0787	-1.1002	-1.1037	-1.1031
1.25	-0.8173	-0.7697	-0.7160	-0.7427	-0.7497	-0.7483
1.50	-0.5967	-0.5667	-0.5727	-0.5644	-0.5624	-0.5619
1.75	-0.4401	-0.4327	-0.4978	-0.4591	-0.4451	-0.4440
2.00	-0.3125	-0.3147	-0.3960	-0.3496	-0.3265	-0.3230

Time	$\bar{x}^8_{AVE}$	$\bar{x}^{32}_{AVE}$	$\bar{x}^4_{SPL}$	$\bar{x}^8_{SPL}$	$\bar{x}^{32}_{SPL}$	$\bar{x}$
0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	-0.1062	-0.2056	-0.2457	-0.2414	-0.2445	-0.2473
0.50	0.1351	0.1245	0.1377	0.1123	0.1104	0.1078
0.75	0.3456	0.4917	0.5717	0.5895	0.5725	0.5663
1.00	0.4120	0.5550	0.6459	0.6175	0.6267	0.6186
1.25	0.3805	0.4232	0.3902	0.4099	0.4204	0.4127
1.50	0.3269	0.2966	0.1737	0.2447	0.2582	0.2474
1.75	0.2988	0.2644	0.2079	0.2420	0.2404	0.2272
2.00	0.3106	0.3115	0.3699	0.3289	0.3230	0.3053

$\bar{J}^N$	2.177	2.401	2.541	2.523	2.523	
CP Time (sec)	31.87	72.39	19.98	49.42	125.27	

ORIGINAL PAGE IS
OF POOR QUALITY

From Table 4 it is evident that the spline approximation for N = 4 is generally better than the averaging approximation for N = 8, at a considerable savings in CP time. A similar comparison may be made between $(\bar{x}^8_{SPL}, \bar{u}^8_{SPL})$ and $(\bar{x}^{32}_{AVE}, \bar{u}^{32}_{AVE})$. One should note however that, for the same value of N, computing solutions via the averaging method is much less costly than by the spline-based scheme, although at a loss in accuracy. Analogous statements may be made about a comparison between the averaging and spline methods for Example 4.3 of [3], which is identical to this example except that the initial function is the constant function $\phi \equiv 10$. It is clear then that the observations made here reflect more than the possibility that a nonconstant ϕ is better approximated by splines than by averaging schemes. The two methods performed equally well when applied to Example 4.2 of the same reference.

§5 Concluding Remarks

The convergence proofs constructed in §3, together with the convergence results obtained in practice lead us to conclude that our method is indeed a reasonable approach to take in the solution of hereditary control problems, performing better than the averaging approximations of [3] in some of the examples we have tried. In addition, the spline-based schemes appear to offer a somewhat inexpensive alternative to existing numerical algorithms in that approximating systems as small as N = 2 or 4 provided satisfactory solutions to many of the control problems we tested.

In future efforts we hope to investigate the possibility of relaxing some of the hypotheses made in §3 (in particular, the global Lipschitz criteria) and to make comparisons between our method and the approach taken in [25], [26], where a full discretization of the FDE control system is examined.

Acknowledgement

The results reported here are a portion of the author's doctoral dissertation written under the supervision of Professor H. T. Banks at Brown University. The author is indebted to Professor Banks for his many valuable comments and suggestions during the course of this work.

References

ORIGINAL PAGE IS
OF POOR QUALITY

[1] E. S. Armstrong and J. S. Tripp, An application of multivariable design techniques to the control of the National Transonic Facility, NASA Technical Paper 1887, Langley Research Center, Hampton, VA, August 1981.

[2] A. V. Balakrishnan, Applied Functional Analysis, Springer-Verlag, New York, 1981.

[3] H. T. Banks, Approximation of nonlinear functional differential equation control systems, J. Optimization Theory Appl., 29 (1979), 383-408.

[4] H. T. Banks, Identification of nonlinear delay systems using spline methods, Proc. International Conf. on Nonlinear Phenomena in the Math Sciences, (Univ. Texas, Arlington, June 16-20, 1980), Academic Press, to appear.

[5] H. T. Banks and J. A. Burns, An abstract framework for approximate solutions to optimal control problems governed by hereditary systems, Proc. International Conf. on Differential Equations, (Univ. So. Calif., Sept. 1974), H. A. Antosiewicz, ed., Academic Press, New York, 1975, 10-25.

[6] H. T. Banks and J. A. Burns, Hereditary control problems: numerical methods based on averaging approximations, SIAM J. Control and Optimization, 16 (1978), 169-208.

[7] H. T. Banks, J. A. Burns and E. M. Cliff, A comparison of numerical methods for identification and optimization problems involving control systems with delays, Brown Univ. LCDS Tech Rep. 79-7, Providence, RI, 1979.

[8] H. T. Banks and P. L. Daniel, Estimation of delays and other parameters in nonlinear functional differential equations, Brown Univ. LCDS Tech. Rep. 82-2 , Providence, RI, 1982, submitted to SIAM J. Control and Optimization.

[9] H. T. Banks and P. L. Daniel, Parameter estimation of nonlinear non-autonomous distributed systems, Proc. 20th IEEE Conf. on Decision and Control, (San Diego, Dec. 16-18, 1981), 228-232.

[10] H. T. Banks and F. Kappel, Spline approximations for functional differential equations, J. Differential Equations, 34 (1979), 496-522.

[11] H. T. Banks and I. G. Rosen, Spline approximations for linear non-autonomous delay systems, ICASE Rep. No. 81-33, Oct. 1981, NASA Langley Research Center, Hampton, VA, to appear in J. Math Anal. Appl.

[12] V. Barbu, Nonlinear Semigroups and Differential Equations in Banach Spaces, Noordhoff, Leyden, 1976.

[13] L. D. Berkovitz, Optimal Control Theory, Springer-Verlag, New York, 1974.

[14] P. L. Daniel, Spline-based approximation methods for the identification and control of nonlinear functional differential equations, Ph.D. Thesis, Brown University, Providence, RI, June, 1981.

ORIGINAL PAGE IS
OF POOR QUALITY

[15] N. Dunford and J. T. Schwartz, Linear Operators, Part I, Interscience, New York, 1958.

[16] G. B. Folland, Introduction to Partial Differential Equations, Princeton University Press, Princeton, New Jersey, 1976.

[17] L. M. Graves, The Theory of Functions of Real Variables, McGraw-Hill, New York, 1956.

[18] J. K. Hale, Ordinary Differential Equations, Interscience, New York, 1969.

[19] F. Kappel, An approximation scheme for delay equations, Proc. International Conf. on Nonlinear Phenomena in the Math Sciences, (Univ. Texas, Arlington, June 16-20, 1980), Academic Press, to appear.

[20] F. Kappel, Spline approximation for autonomous nonlinear functional differential equations, to appear.

[21] F. Kappel and W. Schappacher, Autonomous nonlinear functional differential equations and averaging approximations, Nonlinear Analysis: Theory, Methods and Appl., 2 (1978), 391-422.

[22] E. B. Lee and L. Markus, Foundations of Optimal Control Theory, John Wiley, New York, 1967.

[23] D. G. Luenberger, Optimization by Vector Space Methods, John Wiley, New York, 1969.

[24] L. W. Neustadt, Optimization: A Theory of Necessary Conditions, Princeton University Press, Princeton, New Jersey, 1976.

[25] I. G. Rosen, A discrete approximation framework for hereditary systems, J. Differential Equations, 40 (1981), 377-449.

[26] I. G. Rosen, A discrete approximation framework for hereditary systems, Ph.D. Thesis, Brown University, Providence, RI, June, 1980.

[27] M. H. Schultz, Spline Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1973.

[28] M. H. Schultz and R. S. Varga, L-splines, Numer. Math., 10 (1967), 345-369.

[29] L. Tavernini, One-step methods for the numerical solution of Volterra functional differential equations, SIAM J. Numer. Anal., 8 (1971), 786-795.